INTEGRATED DESIGN OF ULTRADURABLE, LOW CO_2 ALTERNATIVE BINDER SYSTEMS VIA MACHINE LEARNING

Dr. Newell R. Washburn, Carnegie Mellon University

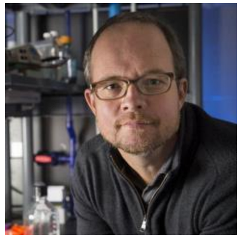
Dr. Kimberly E. Kurtis, Georgia Institute of Technology

Dr. Barnabas Poczos, Carnegie Mellon University

Project Vision

AI-enabled design of sustainable, ultradurable binders

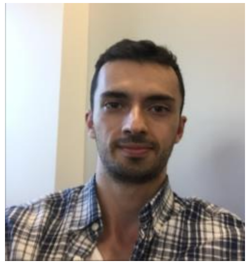
The Team



Newell Washburn (CMU): The Washburn group will characterize, screen, model, and design chemical admixtures and their interactions with binders and concrete, and develop machine learning models.

Barnabas Poczos (CMU): Algorithm design, model assessment, and software development.

Kim Kurtis (Georgia Tech): Characterize and test binder materials, establish relevant engineering models, perform LCA, develop machine learning models, and mix and test cement and concrete.



Ogulcan Canbek (Georgia Tech): Ogulcan is a PhD student at Georgia Institute of Technology whose research focuses on the development of sustainable cement-based materials and cement-admixture interactions through the utilization of statistical modelling with microscale and macroscale characterization techniques

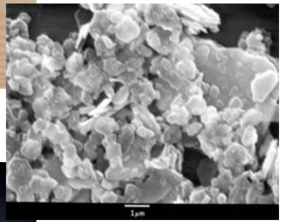
Chris Childs (CMU): Chris completed a PhDin Chemistry at CMU on the study of complex material systems, with primary focus on cementitious systems.

Francesca Lolli (Georgia Tech): Francesca is a post-doctoral associate at Georgia Tech. Her focus is on Life Cycle Assessment and Cost Estimate of C$A and LC3 cement.

Calvin Gang (CMU): Calvin is a PhD student in Chemistry whose research focuses on designing complex materials with machine learning.

Next-Generation Cementitious Materials

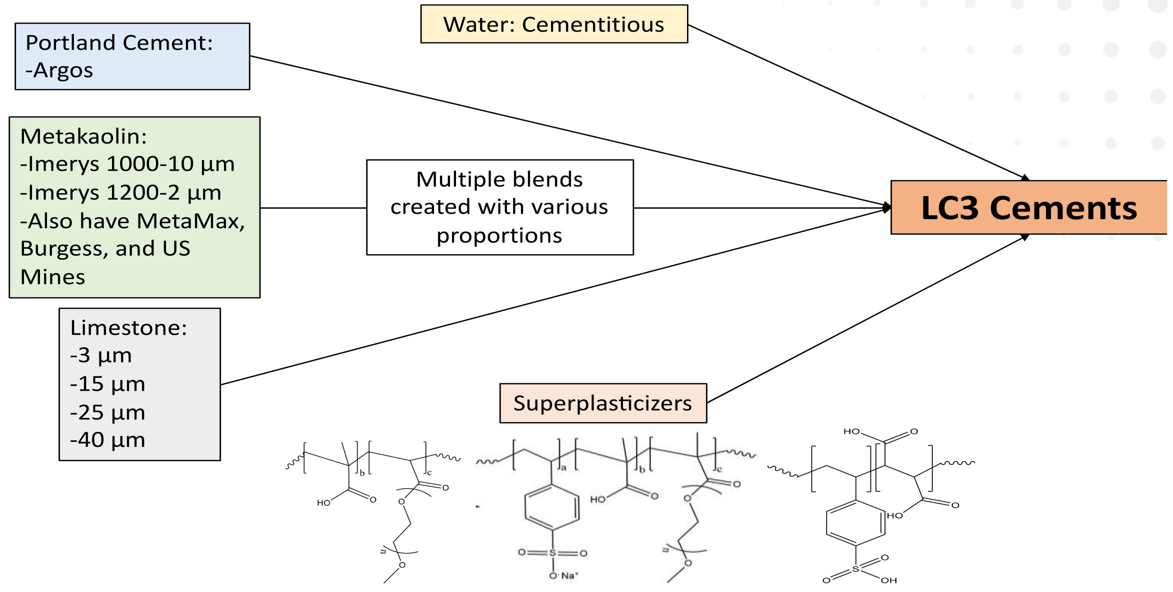
- To minimize embodied CO_2 and maximize performance, cementitious binders will be designed with a diversity of SCMs in complex formulations
  - Limestone, fly ash (spec and off-spec), slag, calcined clays
- Tools for mix design are needed to provide accurate predictions for complex formulations
- In this ARPA-E project, we leverage machine learning to meet sustainability challenges in cementitious infrastructure materials
  - Prototype design tool demonstrated for LC3 with the targets:
    - Half the embodied CO_2
    - Twice the durability
    - Similar handling characteristics and cost as OPC



- **Cost:**
  - **OPC: $110/ton**
  - **LC3: $60-$190/ton** (depending on clay source, LS particle size).

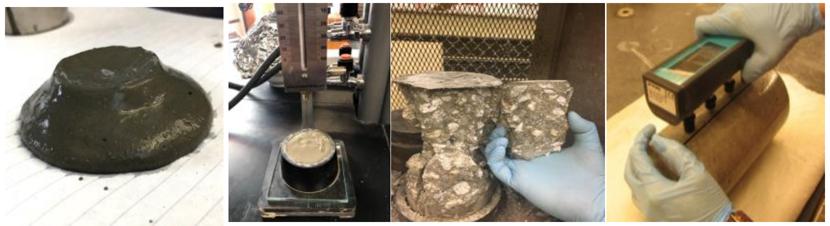
Prototype Design Space

- LC3 has high loading of OPC with calcined clay and limestone
  - Target >50% replacement
- Hypothesis: For some arbitrary combination of materials there is a formulation which will meet design criteria
- For given feedstock sources, develop a tool that specifies:
  - Particle size
  - Limestone and calcined clay loading
  - Limestone:calcined clay ratio
  - Water:binder ratio
  - Superplasticizer



Our Approach: Hierarchical Machine Learning

material properties



workability set time strength durability

latent variables: particle size | packing | pore solution | hydration

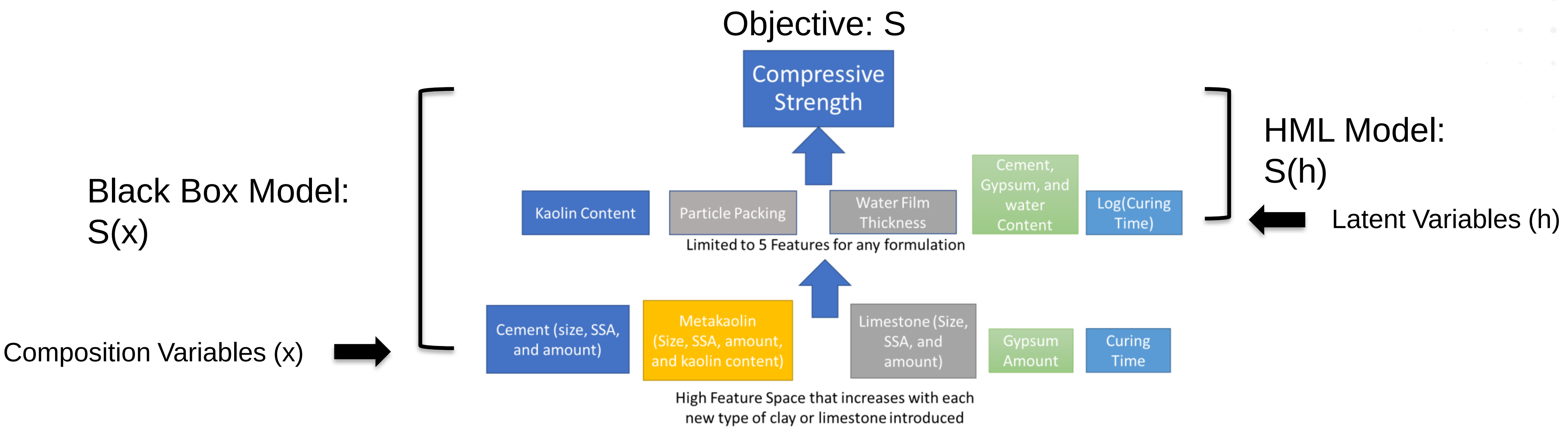
material variables: OPC | calcined clay | limestone | water | admixtures

statistical learning

physical modeling

- HML integrates physical modeling in a framework of statistical learning, allowing accurate predictions from small datasets and facile transfer learning to different feedstocks
- Models for workability and strength were separately developed

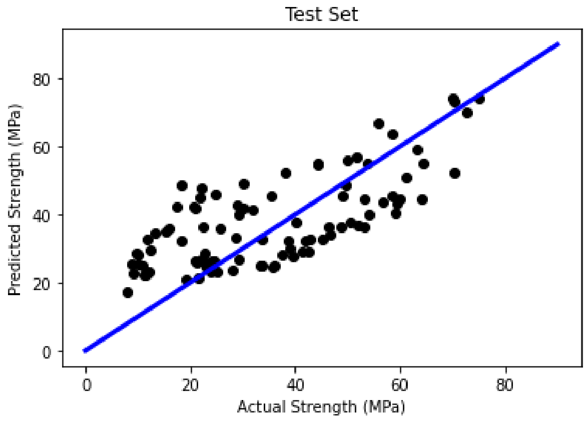
Methodology for Predicting Binder Strength: Black Box vs. HML Models



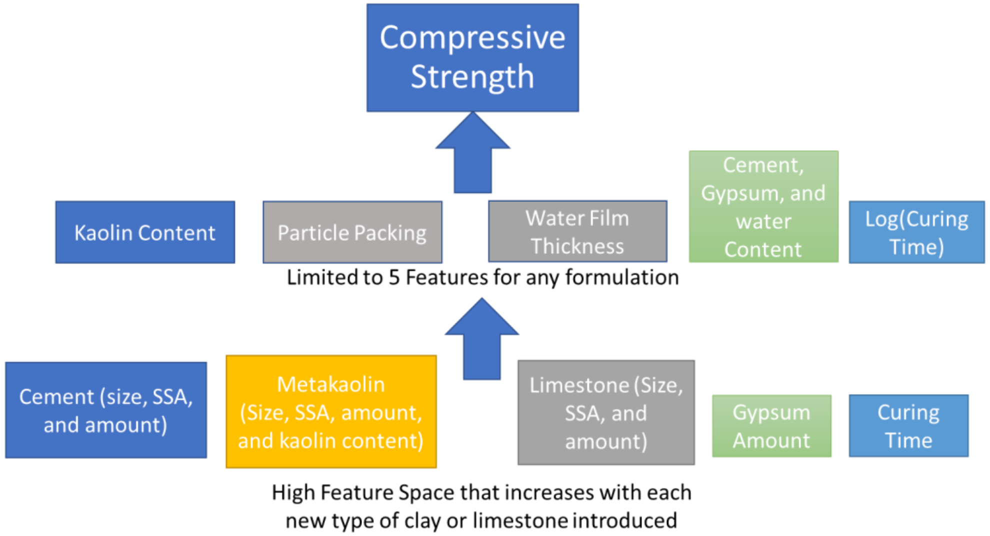
- In a “black box” approach, the relationship between strength and composition is modeled by measuring the strength of many compositions and fitting the resulting curves with machine learning tools
- In an HML approach, strength is modeled by estimating latent variables, such as particle packing and pozzolanic activity
  - Provides functional relationships that connect the predicted values of strength

HML More Accurate than Black Box Model

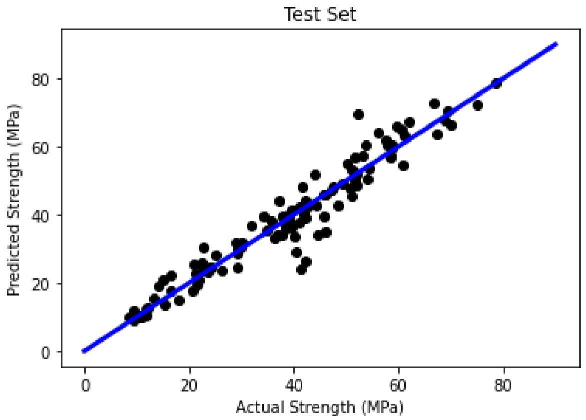
Black Box Model



Train RMSE: 11.29 MPa
Test RMSE: 12.81 MPa
Train R^2: 0.60
Test R^2 : 0.49



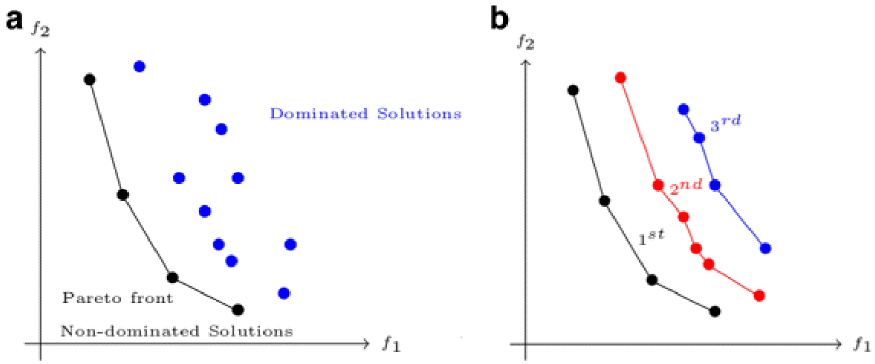
HML Model

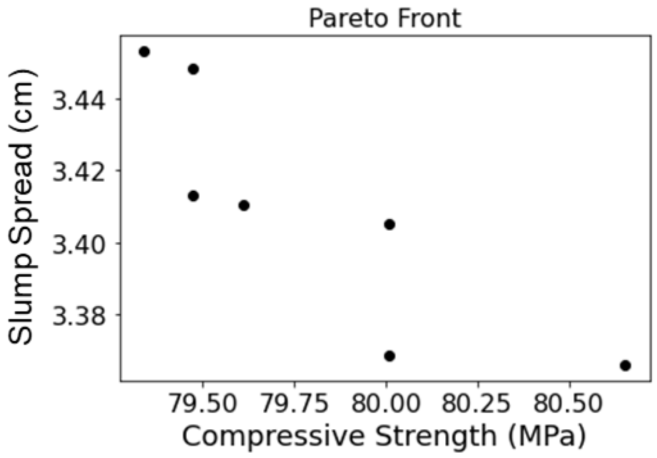


Train RMSE: 2.59 MPa
Test RMSE: 4.65 MPa
Train R^2: 0.98
Test R^2 : 0.92

Results for LC3 multi-objective optimization: Simultaneous Prediction of Workability and Strength

- The Pareto front was determined to maximize mini-slump spread and strength subject to constraints:
- Cost < $140/ton
- CO_2 < 450kg/ton





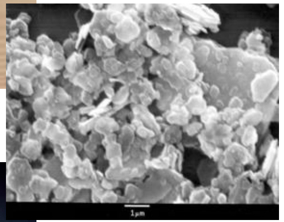
Cement	MK1000	MK1200	LS3	LS15	LS25	LS40	Metamax	Gypsum	w/binder
0.55	0.00	0.06	0.09	0.00	0.00	0.19	0.09	0.02	0.40

SP%	Predicted PAT (cm)	Measured PAT (cm)	ARPA-E Goal PAT (cm)	Set Time (min)	ARPA-E Goal Set Time (min)
0.25%	5.8cm	4.6cm	4.0cm	130min	60min

Cure Time	Predicted Strength (MPa)	Measured Strength (MPa)	ARPA-E Goal Strength (MPa)
7 days	66 MPa	46 MPa	21 MPa
28 days	79 MPa	46 MPa	28 MPa

AI-enabled design of LC3

- LC3 is the leading candidate to be a more sustainable replacement for OPC
- Hypothesis: Mix design for LC3 is a complex formulation problem
  - For some arbitrary combination of portland cement, limestone, and calcined clay, there is a formulation that will meet performance criteria
  - A machine-learning design tool is developed for optimizing LC3 formulations from arbitrary feedstocks



Project Objective: Design tools based on machine learning

- In general, a machine learning algorithm takes a compositional space and creates a relation to an output.
- Benchmark study: 28-day compressive strength models for concrete of 706 samples from laboratory-produced samples (Yeh et al.) and 9994 samples from a range of various concrete production sites (VIP). R^2 values ranged from 0.54 to 0.86.

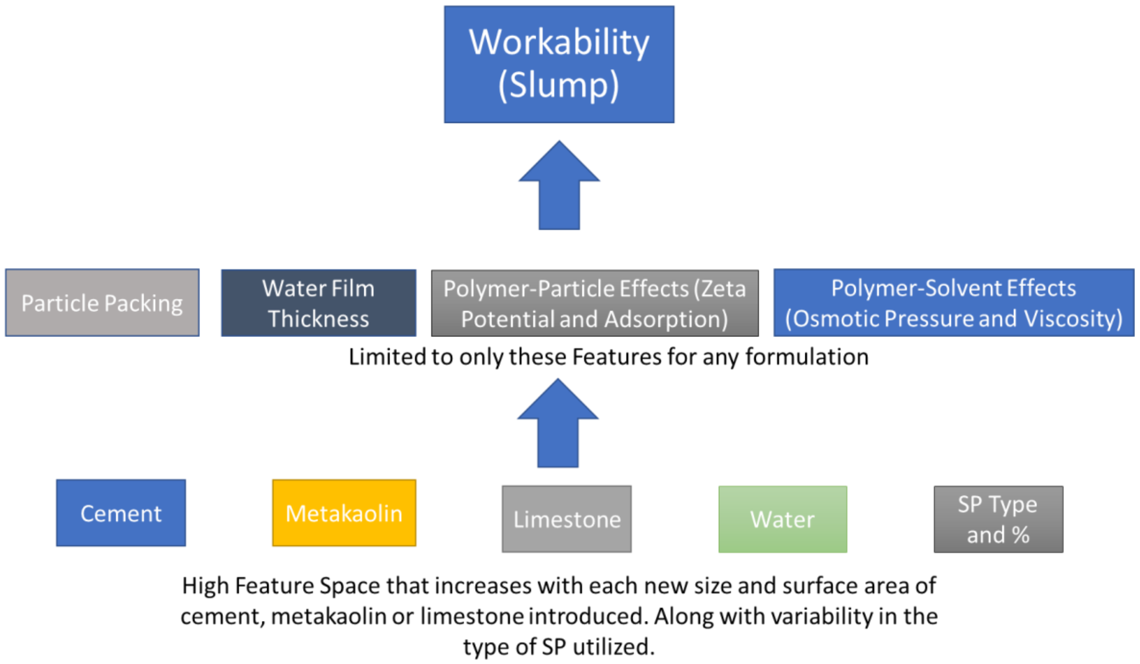
Mix parameter
Cementitious material content
w/c
Fly ash content
Coarse aggregate content
Fine aggregate content
Fly ash/total cementitious material ratio
Total volume fraction of aggregates
Coarse/fine aggregate ratio

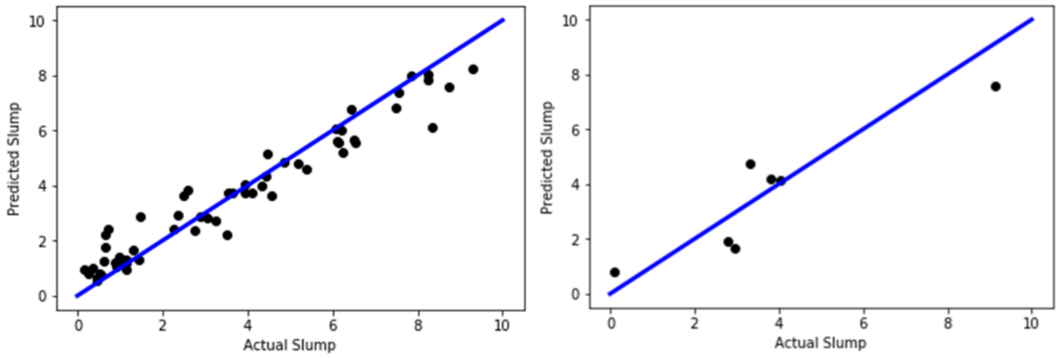
Model	Yeh et al. dataset			VIP data set		
	RMSE (MPa)	R^2	MAPE (%)	RMSE (MPa)	R^2	MAPE (%)
Linear regression	8.8	0.66	22	5.0	0.49	10
Neural network	6.3	0.82	14	4.8	0.54	9
Random forest	5.7	0.86	14	4.4	0.60	9
Boosted tree	5.8	0.85	13	4.5	0.59	9
SVM	6.4	0.83	15	4.5	0.59	9

Young, B. A., Hall, A., Pilon, L., Gupta, P. & Sant, G. Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods. *Cem. Concr. Res.* **115**, 379–388 (2019).

Model for LC3 workability

- A random forest was utilized for the prediction of slump based on 66 various LC3 formulations and PCE architectures. Utilizing latent variables performs better than composition variables for training data and generalizing to a test set.

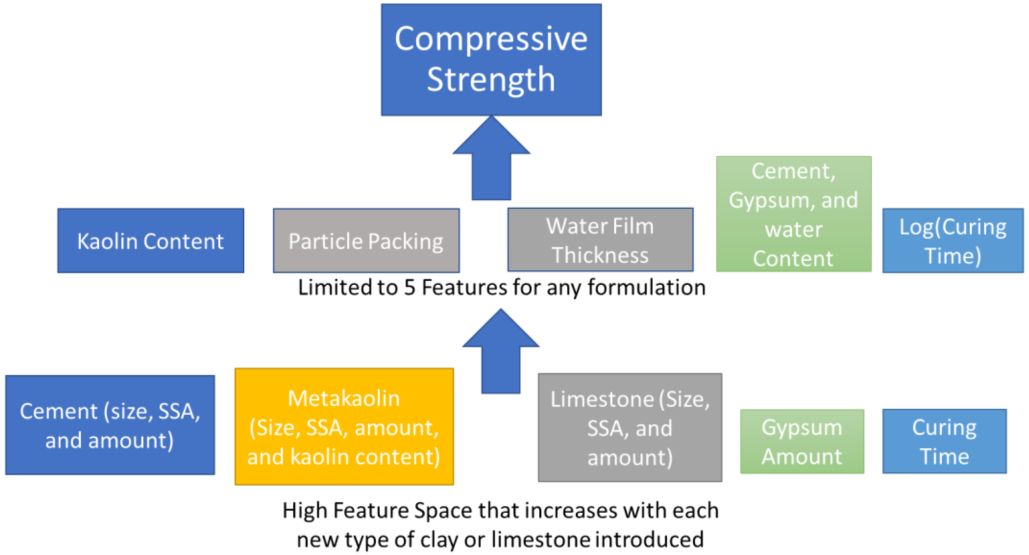


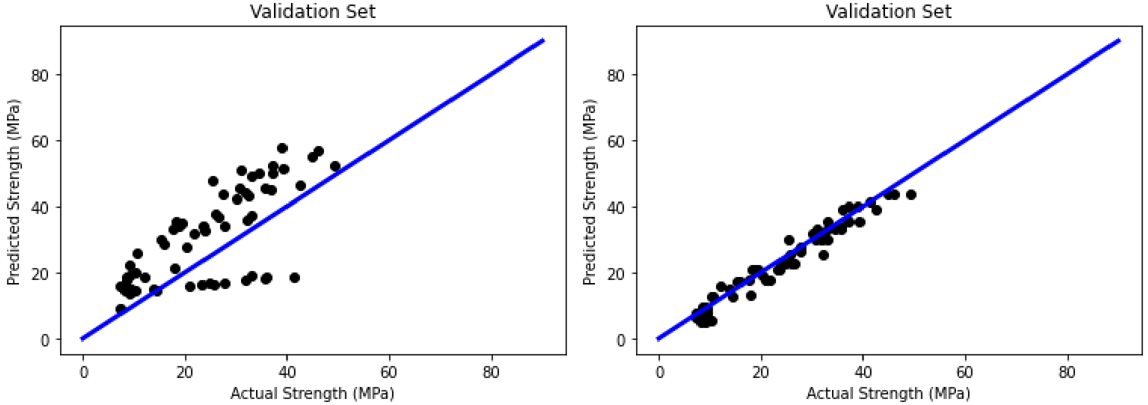


	R^2	MSE	RMSE (cm)	Expected Error %
Training Set Bottom	0.82	1.3	1.14 cm	8.7%
Training Set Middle	0.93	0.52	0.72 cm	7.7%
Test Set Bottom	0.71	1.82	1.35 cm	14.5%
Test Set Middle	0.83	1.07	1.03 cm	11.0%

Modeling LC3 strength

- A gaussian process regression was utilized for the prediction of strength based on 96 LC3 formulations, with a total of 435 datapoints for various days of testing. Utilizing a middle layer performs better in terms of generalizing to a validation set which was tested with different types of materials and higher w/cm ratios than any point in the training set.





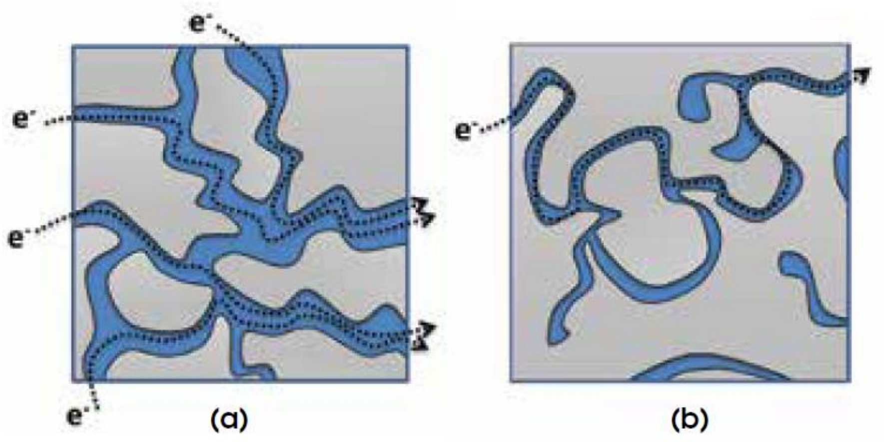
	R^2	RMSE (MPa)
Validation Set Bottom	-0.02	11.74 MPa
Validation Set Middle	0.95	2.54 MPa

Modeling LC3 durability

- Surface resistivity (AASHTO T 358-15) is a rapid indication of concrete resistance to the penetration of chloride ions
- High correlation with chloride exposure tests such as ASTM C1202 and ASTM C1556
- Development of surface resistivity over time can show microstructural development during cement hydration
- An indirect method to assess pozzolanicity of supplementary cementitious materials

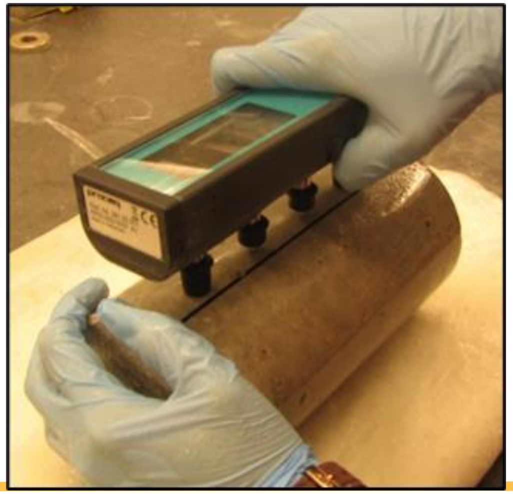
High permeability

Low resistivity



Low permeability

High resistivity

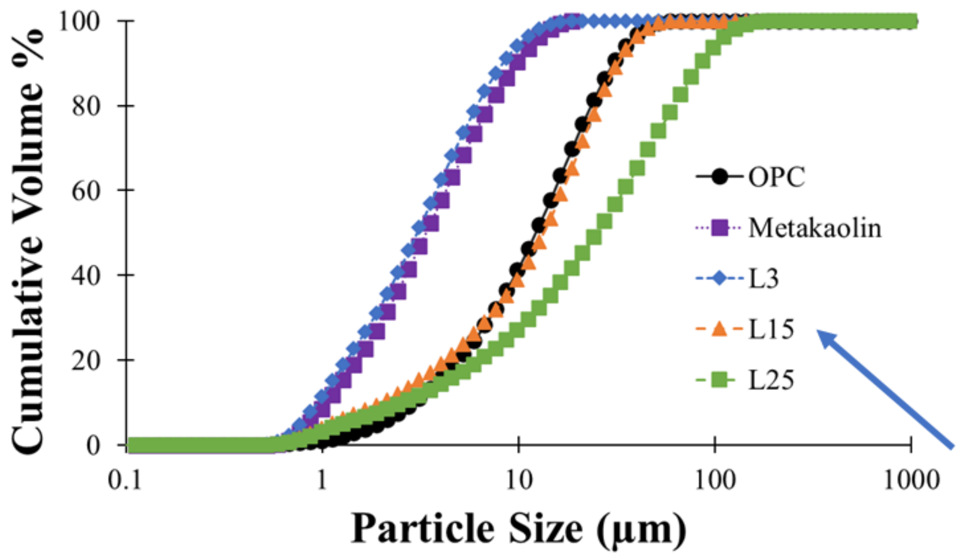


A schematic from Nadelman and Kurtis (2014) showing electron flow in concrete

Wenner probe

Training set

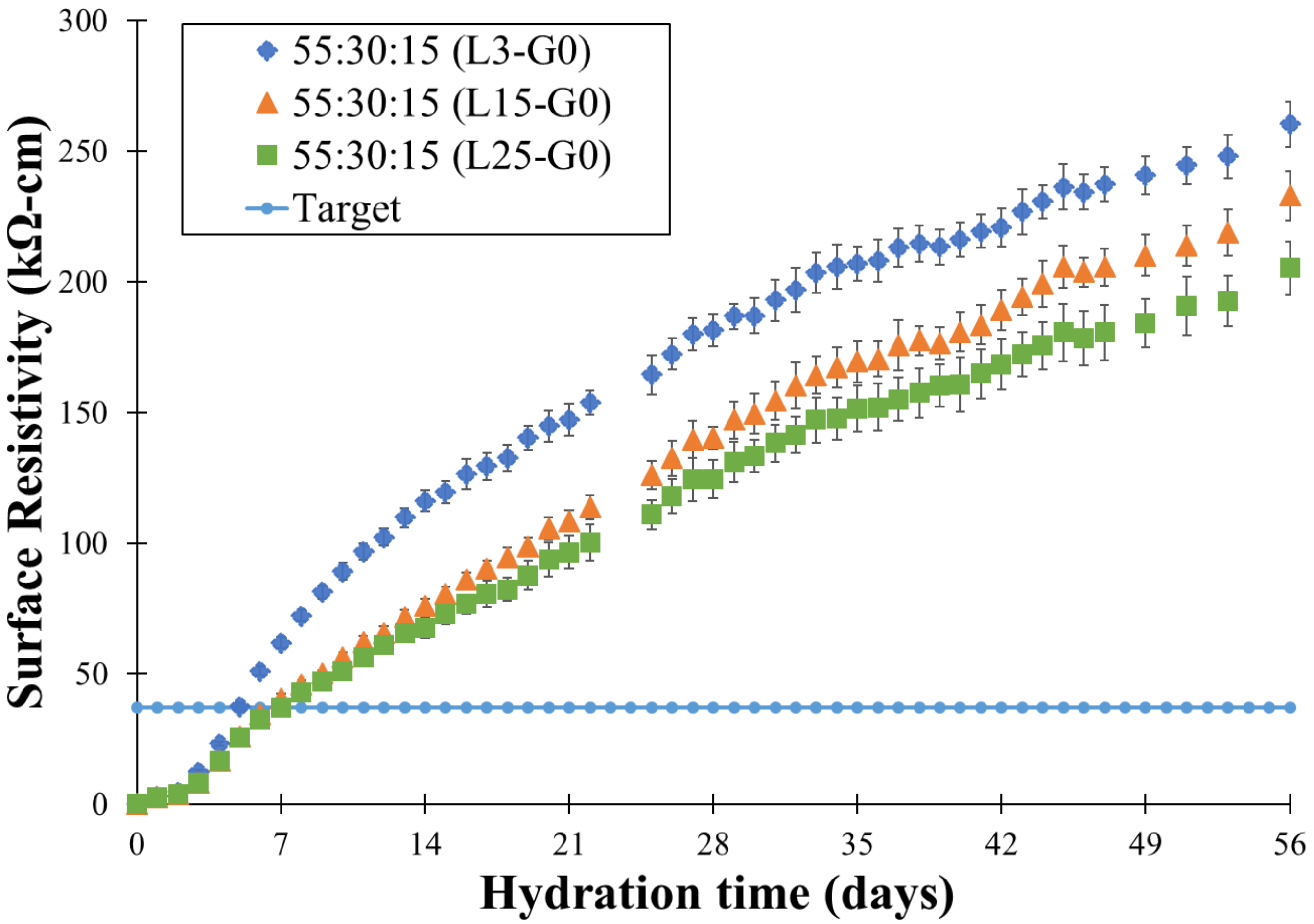
- Surface resistivity of 11 concrete mixtures including 7 different mixture proportions, 3 different limestone particle sizes (L3, L15, L25) and 3 different gypsum addition levels (0%, 2%, and 5% by mass of solid) were measured up to 56 days of hydration
- Water-to-cementitious (w/c) ratio was kept constant at 0.4, and up to 1% (wt.% of solid) superplasticizer was added



Limestone source with a median particle diameter (D_{50}) of 15 µm is abbreviated as L15

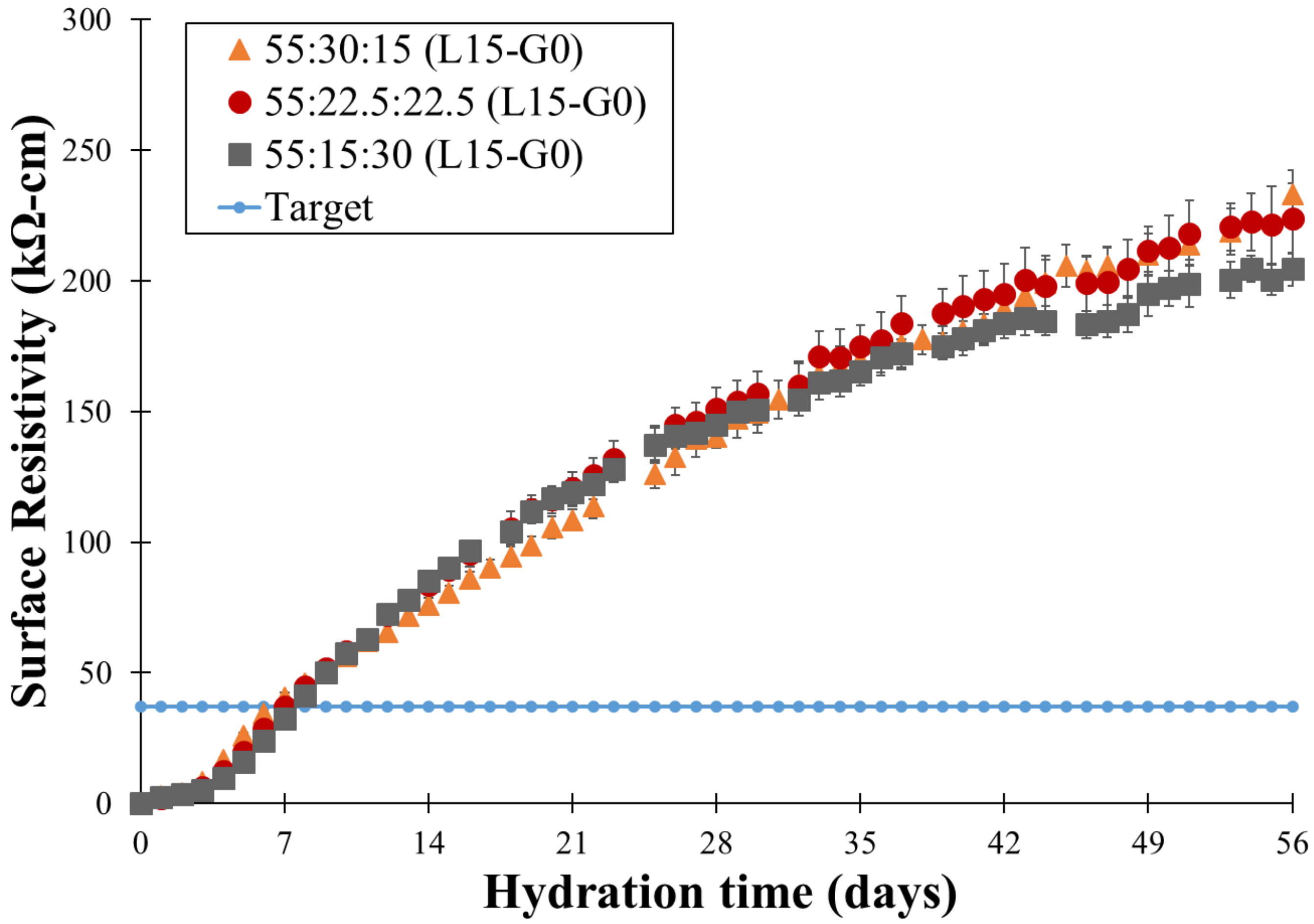
Particle size distribution of materials used

Increasing limestone fineness results in higher resistivity



- Filler effect imparted by L3 enhances the surface resistivity
- By 8 days of hydration, the target surface resistivity of 37 kΩ-cm is achieved for all mixes

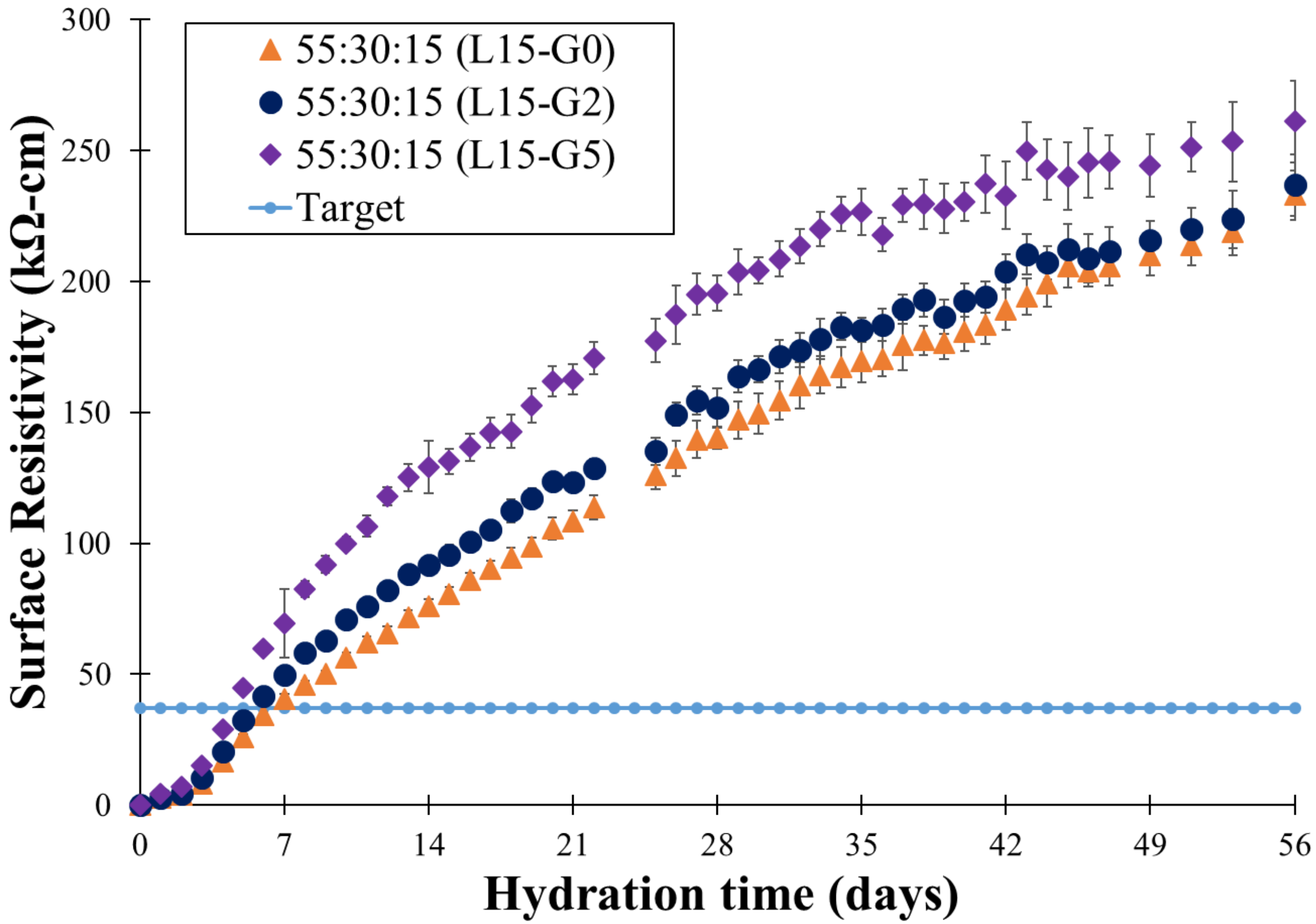
Complex relationship between metakaolin content and surface resistivity



- Direct correlation between metakaolin content and surface resistivity only up to 8 days
  - Early pore refinement depending on metakaolin

- Highest metakaolin including mix -55:30:15 (L15-G0)- has slightly lower resistivity after 8 days

Gypsum addition enhances the surface resistivity of LC^3



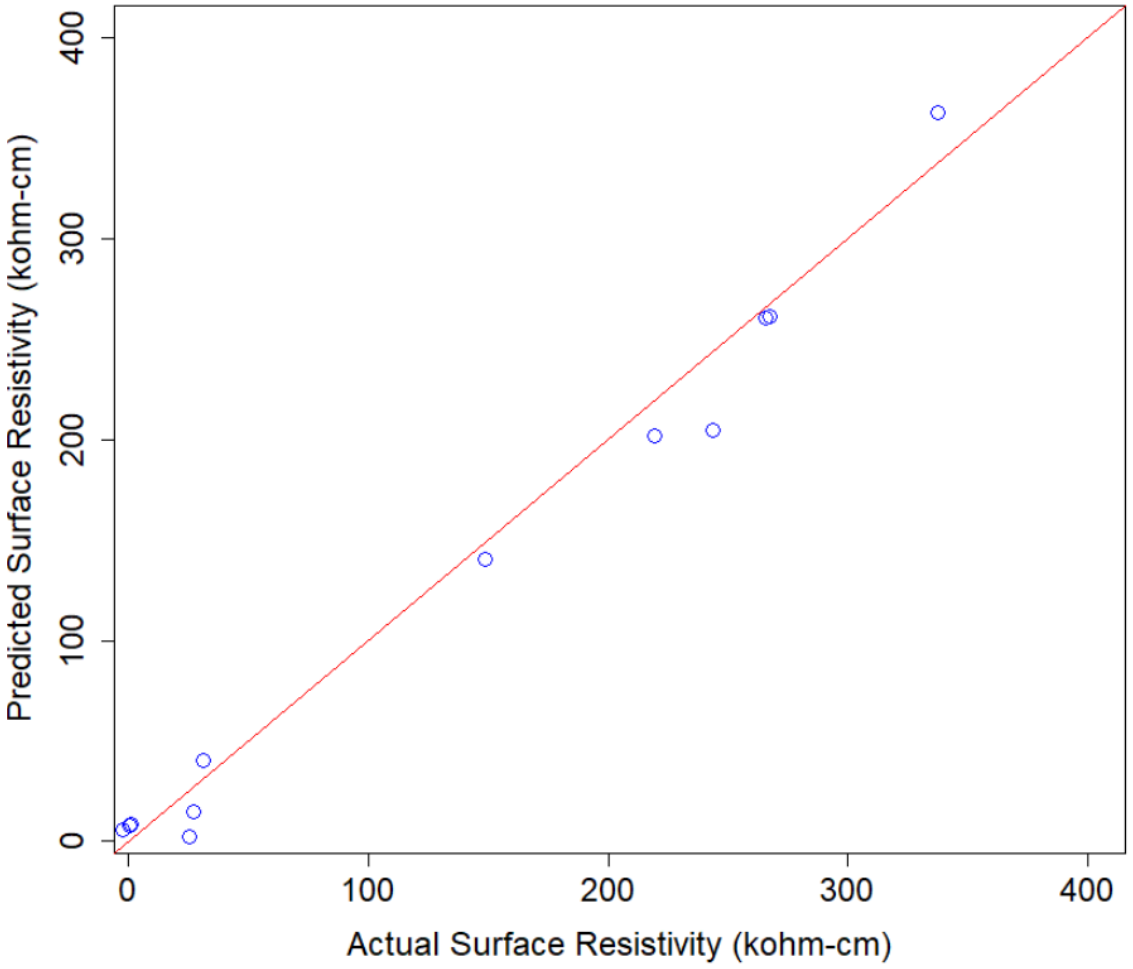
- Ettringite formation is enhanced with extra gypsum, resulting in higher surface resistivities
- By 7 days of hydration, all mixes achieved the target surface resistivity of 37 kΩ-cm

PREDICTING SURFACE RESISTIVITY OF LC^3 CONCRETE

- All 11 mixes LC3 were included in modeling
- The dataset was split to 75% training and 25% testing sets with cross validation
- Stepwise, LASSO and Support Vector Machine (SVM) algorithms were implemented with predictor (X) variables selected as follows:
  - OPC content (wt.%)
  - Metakaolin/Limestone mass ratio
  - Limestone median particle diameter
  - SO_3 content in cementitious mix (%)
  - Al_2O_3/SO_3 in cementitious mix
  - Log(time)
- For SVM, Radial kernel was used, and Cost and gamma hyperparameters were tuned through grid search

SUPPORT VECTOR MACHINE ACCURATELY PREDICTS SURFACE RESISTIVITY

Model	RMSE (kohm-cm)
Stepwise	50.73
LASSO	45.32
SVM (Radial Kernel)	**18.37**



SVM model fitting graphic: Actual vs. Predicted

Toward Commercialization

- The algorithms developed in this ARPA-E project will be further developed by Ansatz AI
- Ansatz AI is a chemicals/materials informatics company founded by Profs. Washburn and Poczos
  - Accelerated design and optimization of complex systems
- Team member Chris Childs joined Ansatz AI and will lead this effort in partnership with the cement and concrete industry

Conclusions and Acknowledgments

- Machine learning models of LC3 based on latent variables are accurate and effective for mix design and multi-objective optimization
  - Work continues on identifying latent variables from experiments and theory for more accurate machine learning models
- Rapid virtual screening of chemical admixtures can be performed using cheminformatics and machine learning
  - Rapid transfer learning to new binder systems

- Acknowledgments:
- Dr. Joseph King (ARPA-E)
- Prof. Karen Scrivener (EPFL)
- Dr. Franco Zunino (EPFL)